

Jakub Podolak

✉ jakub.podolak.43@gmail.com • 🔗 linkedin.com/in/jakub-podolak
github.com/jjp241/

Experience

Data Scientist

dotData, remote

Warsaw (Remotely)

Jul 2023 – Current

- Key decision-making position in research and development of the automated **feature discovery** platform.
- Implemented **gradient boosting** inspired methods for automated feature discovery. Prototyped new features with **OpenAI models**, utilising **vector embeddings** and **in-context learning** techniques.
- Delivered 3 new prototypical features to product, currently in trial, with one already presented to public.

ML Research Assistant

Ideas NCBR, Poland

Warsaw

Apr 2023 – Jul 2023

- Implemented **early exit methods** in Computer Vision models in **PyTorch**, based on research papers.
- Benchmarked efficiency of **Dynamic Architectures** in ResNets and **Vision Transformers**.
- Performed experiments on a distributed GPU cluster with slurm workloader, with focus on **accuracy-performance tradeoff** in Machine Learning.
- Internship continuing to publication at conference.

Intern Researcher

University of Warsaw, Poland

Warsaw

Oct 2022 – Current

- Researched usage of **open-source LLMs** and **OpenAI Models** to automatically detect disinformation and hate speech, and provide fact-checked responses.
- Statistical results testing - **A/B testing**, **bootstrapping** and statistical significance measuring.
- Utilised **Huggingface transformers** library to quickly prototype and evaluate models.
- Functional programming in **OCaml** and **Haskell**
- Bachelor Thesis continuing work towards submitting a conference publication.

Junior Machine Learning Specialist

MIM-Solutions, Poland

Warsaw

Apr 2021 – Jan 2023

- Developed **LGBM Model** to predict the conversion rate of web ads, with the use of **Robust Neural Networks** and **PyTorch** to extract ads' image features. Using those predictions, we **increased profit per click by 90%**.
- Trained, validated and measured performance of **SVC models** for malicious website classification. Contributed to text-preprocessing pipeline, and data gathering.
- Created a computer vision pipeline for low-quality medical documents digitization with use of **OpenCV**, **custom Hough transform** and **Google OCR**.

Intern/Junior React Developer

Spinney, Poland

Krakow

Jun 2020 – Nov 2020

- Developed **React.js/react-native** components and implemented updates in mobile/web game with over **5 million downloads**.

Robotics and Python Teacher

Primary and High Schools, Poland

Krakow

Jun 2018 – March 2021

- Taught High School students programming in **Python** and created together a computer game with the PyGame library.

Education

MSc Artificial Intelligence

University Of Amsterdam, Netherlands

Amsterdam

Sep 2023 – Jun 2025

- Classic **Machine Learning** course - **Bayesian models**, SVM, Optimizers and neural networks.
- **Computer Vision**, **Deep Learning** courses - **Numpy** and **pyTorch**.

BSc Computer Science

University Of Warsaw, Poland

Warsaw

Oct 2020 – Jun 2023

- o Graduated from Algorithmics and Data Structures in C, C++ **top of class**.
- o PostgreSQL, **OOP in Java**, OCaml, **Statistical Data Analysis**.
- o **Deep learning, NLP course** - Gradient Descent and optimizers, **Computer Vision** and other DNN techniques .
- o **Microservices course** - Building a distributed stock simulation tool with **Docker, Kafka Streams, RabbitMQ**.
- o **English - C1 level** (CAE Cambridge Assesment), Polish - native.

Organisations

ML in PL Association Member

Warsaw

ML in PL, Poland

Dec 2022 – Present

- o Reached out to **Machine Learning researchers/ popularizers** and invited them to panel talk during **ML in PL** conference - one of the best ML conferences in Poland. Took care of the guests and kept them updated.
- o Prepared questions for a panel talk and helping with onsite organisation

SocialAI

Warsaw

SocialAI, Poland

Apr 2022 – Jun 2023

- o Utilised pre-trained polish **BERT** model to predict text sentiment in psychological research. Used **Markov Chains** and **convolution** to model emotion changes in time.

Recognitions

- o University representant at **CERC 2022 - Central European ICPC Contest** - biggest team algorithmic competition in this EU region.
- o **1st Place in Goldman Sachs** 2022 Hackathon in Warsaw.
- o 2nd place in BEST Hacking League Hackathon - **created the AI chatbot app with custom knowledge** based on OpenAI API.
- o Finalist of **Polish Olympiad in Informatics**.
- o 3rd place in Polish CanSat competition where we built and coded our can-sized probe. Organized by ESA (European Space Agency).
- o Multiple Codeforces competitions and algorithmic camps.

Courses and participations

- o **NVIDIA** Fundamentals of Deep Learning certificate.
- o Speaker at **Data Science Summit** ML Edition 2021 - talked about robust neural networks and feature extraction.
- o Speaker at **PyKonik** local python meetup in Krakow.
- o **MIT Intro to Deep Learning** - MIT 6.S191 online course containing **RNNs, LSTMs, GANs, RL**.
- o Participated in **EuroPython 2019/2022** in Basel and Dublin - **NLP, ML and DS** techniques.
- o **Scrum/Agile** course in Ringier Axel Springer Poland.
- o Speaker at Seminar in MIM Solutions company - told about **Meaningful latent image representations** using **ResNet** classifier.

Hobbies

- o Creating electronic music - some releases on Spotify, cooking, 3D-printing